

# SEO SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 06:14:15 ; Search time 16680 Seconds  
(without alignments)  
11529.559 Million cell updates/sec

Title: US-09-753-143-72  
Perfect score: 4437  
Sequence: 1 GCGCGGCGGCCGTGGTTGCG.....AAAAAAAAAAAAAAAAAAAAA 4437

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%				Query		ID	Description
	No.	Score	Match	Length	DB	ID		
	1	4437	100.0	4437	6	AR049083	USPN 5824501	AR049083 Sequence
	2	4435.4	100.0	4437	6	AR211947	" 6399378	AR211947 Sequence
	3	4435.4	100.0	4437	6	AX335489	Wo 01/94629	AX335489 Sequence
	4	4435.4	100.0	4437	6	AX695780	Wo 03/08583	AX695780 Sequence
	5	4435.4	100.0	4437	9	HSU39817	GenBank	U39817 Human Bloom
	6	4252.4	95.8	4254	6	AX695781		AX695781 Sequence
	7	2925.8	65.9	4498	6	AX695777		AX695777 Sequence
	8	2925.8	65.9	4498	10	MMBLM	GenBank	Z98263 Mus musculu
	9	2888	65.1	4763	10	AB008674	"	AB008674 Mus muscu
	10	2882.6	65.0	4251	6	AX695778		AX695778 Sequence
	11	1204.2	27.1	3581	5	AB040747	GB	AB040747 Gallus ga
	12	1082.8	24.4	4124	5	AF307841		AF307841 Xenopus l
c	13	704	15.9	95565	9	AC000379	GB	AC000379 Human Chr 1997
	14	704	15.9	96594	6	AX695779		AX695779 Sequence
	15	704	15.9	99500	6	AR211954		AR211954 Sequence
	16	704	15.9	147854	9	AC021422		AC021422 Homo sapi
c	17	704	15.9	157248	9	AC002312		AC002312 Human Chr
c	18	670.4	15.1	100259	9	AC005800		AC005800 Homo sapi
	19	670.4	15.1	142201	9	AC124248		AC124248 Homo sapi
c	20	558.8	12.6	64525	2	AC124238		AC124238 Homo sapi
	21	534	12.0	630	6	AR211955		AR211955 Sequence
	22	460.8	10.4	514	9	BC062697		BC062697 Homo sapi
c	23	450.6	10.2	198446	2	AC110907		AC110907 Mus muscu
c	24	441.2	9.9	94175	2	AC136849		AC136849 Rattus no
	25	425.6	9.6	96593	6	AX695776		AX695776 Sequence
	26	359.6	8.1	260924	2	AC095740		AC095740 Rattus no
	27	335.6	7.6	4835	3	DMU92536		U92536 Drosophila
	28	305.4	6.9	3172	3	AY095296		AY095296 Caenorhab
	29	299.2	6.7	2060	8	ATH404470		AJ404470 Arabidops
	30	269.2	6.1	3567	8	BT010133		BT010133 Arabidops
	31	269.2	6.1	3891	8	AY120761		AY120761 Arabidops
	32	269.2	6.1	3916	8	ATH404473		AJ404473 Arabidops
	33	268	6.0	5161	8	SPDNAHEL		Y09426 S.pombe hus

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 06:09:25 ; Search time 1584 Seconds  
(without alignments)  
11899.782 Million cell updates/sec

Title: US-09-753-143-72  
Perfect score: 4437  
Sequence: 1 GCGCGGCGGCCGTGGTTGCG.....AAAAAAAAAAAAAAAAAAAAA 4437  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4435.4	100.0	4437	6	ABL67661 <i>WO 01/194629</i> Abl67661 Oesophagu
2	4435.4	100.0	4437	6	AAD41733 <i>USPN 6399378</i> Aad41733 Human REC
3	4435.4	100.0	4437	8	ADA02889 <i>WO 03/057146</i> Ada02889 Human BLM
4	4435.4	100.0	4437	9	ADB72627 <i>WO 03/05853</i> Adb72627 Human BLM
5	4435.4	100.0	4437	9	ADC85368 <i>WO 03/045230</i> Adc85368 Human Blm
6	4432.2	99.9	4437	2	AAT67013 <i>WO 97/17979</i> <del>Aat67013</del> Bloom syn
7	4430.6	99.9	4437	2	AAT93390 Aat93390 Bloom's s
8	4430.6	99.9	4437	2	AAT93392 Aat93392 Bloom's s
9	4430.6	99.9	4437	2	AAT93395 Aat93395 Bloom's s
10	4430.6	99.9	4437	2	AAT93394 Aat93394 Bloom's s
11	4421.2	99.6	4438	2	AAT93391 Aat93391 Bloom's s
12	4416.2	99.5	4434	2	AAT93389 Aat93389 Bloom's s
13	4414.8	99.5	4438	2	AAT93393 Aat93393 Bloom's s
14	4252.4	95.8	4254	8	ADA02890 <i>WO 02/057146</i> Ada02890 Human BLM
15	4252.4	95.8	4254	9	ADB72628 Adb72628 Human BLM
16	4252.4	95.8	4254	9	ADC85369 Adc85369 Human Blm
17	2925.8	65.9	4498	8	ADA02886 Ada02886 Mouse Blm

*App.*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 00:56:19 ; Search time 289 Seconds  
(without alignments)  
8520.128 Million cell updates/sec

Title: US-09-753-143-72  
Perfect score: 4437  
Sequence: 1 GCGCGGCGGCCGTGGTTGCG.....AAAAAAAAAAAAAAAAAAAAA 4437

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%				Description
	No.	Score	Query	Match	Length	ID	
c	1	4437	100.0	4437	1	US-08-559-303B-72	Sequence 72, Appl
	2	4437	100.0	4437	3	US-09-175-828-72	Sequence 72, Appl
	3	4435.4	100.0	4437	4	US-09-798-096-3	Sequence 3, Appli
	4	704	15.9	99500	4	US-09-798-096-10	Sequence 10, Appl
	5	534	12.0	630	4	US-09-798-096-11	Sequence 11, Appl
	6	159.8	3.6	1926	4	US-09-543-681A-187	Sequence 187, App
	7	129	2.9	1830121	4	US-09-557-884-1	Sequence 1, Appli
	8	129	2.9	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	9	123.6	2.8	1851	4	US-09-107-532A-3604	Sequence 3604, Ap
	10	120.6	2.7	1914	4	US-09-134-001C-2821	Sequence 2821, Ap
	11	119.2	2.7	2004	4	US-08-956-171E-269	Sequence 269, App
	12	113.6	2.6	1860	4	US-09-489-039A-904	Sequence 904, App

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 05:30:25 ; Search time 1843 Seconds  
(without alignments)  
11804.250 Million cell updates/sec

Title: US-09-753-143-72  
Perfect score: 4437  
Sequence: 1 GCGCGGCGGCCGTGGTTGCG.....AAAAAAAAAAAAAAAAAAAAA 4437

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	% Query					Description
	No.	Score	Match Length	DB	ID	
	1	4437	100.0	4437	9	US-09-753-143-72
	2	4435.4	100.0	4437	9	US-09-962-832-112
	3	4435.4	100.0	4437	12	US-09-997-722-155
	4	4435.4	100.0	4437	13	US-10-342-887-420
	5	4435.4	100.0	4437	13	US-10-172-118-420
	6	4313.4	97.2	4544	16	US-10-062-674-1885
	7	4252.4	95.8	4254	12	US-09-997-722-156
	8	2925.8	65.9	4498	12	US-09-997-722-152
	9	2882.6	65.0	4251	12	US-09-997-722-153
	10	704	15.9	96594	12	US-09-997-722-154
c	11	701	15.8	701	15	US-10-029-386-20899
c	12	581.4	13.1	607	9	US-09-867-701-10217
	13	425.6	9.6	96593	12	US-09-997-722-151
	14	380	8.6	509	16	US-10-062-674-1166
c	15	306	6.9	306	9	US-09-864-761-22841
	16	273.4	6.2	2150	13	US-10-425-114-19444
	17	254.2	5.7	1584	13	US-10-425-114-3024
	18	250.2	5.6	2925	10	US-09-793-807-3
	19	246.4	5.6	1952	17	US-10-437-963-38862
	20	245.4	5.5	567	10	US-09-918-995-28609
	21	241	5.4	2505	13	US-10-425-114-470
c	22	206.6	4.7	365	9	US-09-864-761-5348
c	23	206.6	4.7	365	9	US-09-864-761-14448
	24	204.6	4.6	5868	15	US-10-293-504-2
c	25	179.8	4.1	558	15	US-10-029-386-7199
	26	172	3.9	6916	15	US-10-293-504-1
	27	166.2	3.7	1172	13	US-10-282-122A-25257
	28	161.4	3.6	1800	13	US-10-282-122A-33030
	29	160.6	3.6	2160	13	US-10-282-122A-12367
	30	154.8	3.5	1340	13	US-10-424-599-78917
	31	154.4	3.5	1164	13	US-10-424-599-35459
	32	148.2	3.3	1863	13	US-10-282-122A-40794
	33	143.8	3.2	684707	16	US-10-398-221-9
c	34	143.8	3.2	3011208	16	US-10-398-221-2058
	35	143.2	3.2	1803	13	US-10-282-122A-12144
	36	142.6	3.2	2481	13	US-10-282-122A-16967
	37	142.4	3.2	2127	13	US-10-282-122A-16639
	38	139.6	3.1	2145	13	US-10-282-122A-15954
	39	137.2	3.1	1833	13	US-10-282-122A-41841
	40	137	3.1	2301	13	US-10-282-122A-29918
	41	136.6	3.1	1860	9	US-09-815-242-6988
	42	136.6	3.1	1860	13	US-10-282-122A-22087
c	43	135.4	3.1	2731748	17	US-10-297-465A-1
	44	134.6	3.0	1848	13	US-10-282-122A-38789
	45	133	3.0	1830	13	US-10-282-122A-39599

## ALIGNMENTS

RESULT 1

US-09-753-143-72

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2004, 09:58:50 ; Search time 10409 Seconds  
(without alignments)  
12729.224 Million cell updates/sec

Title: US-09-753-143-72  
Perfect score: 4437  
Sequence: 1 GCGCGGCGGCCGTGGTTGCG.....AAAAAAAAAAAAAAAAAAAAA 4437

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
c	1	4364.2	98.4	4493	11	BC034480 (2007)	BC034480 Homo sapi
	2	947.8	21.4	974	9	AL556823	AL556823 AL556823
	3	942	21.2	1019	9	AL556853	AL556853 AL556853
c	4	810.4	18.3	868	12	BM451903	BM451903 AGENCOURT
	5	803	18.1	964	12	BG199179	BG199179 RST18458
	6	781.8	17.6	849	13	BX434842	BX434842 BX434842
c	7	776.4	17.5	874	13	BX451969	BX451969 BX451969
	8	726.6	16.4	882	12	BI091601	BI091601 602859024
	9	711	16.0	772	12	BG772975	BG772975 602721230
c	10	709.8	16.0	873	13	BX451970	BX451970 BX451970
	11	697	15.7	865	13	BQ230262	BQ230262 AGENCOURT
	12	695.4	15.7	697	12	BI091772	BI091772 602858823
c	13	683	15.4	784	12	BM041661	BM041661 603614723
	14	680.2	15.3	782	14	CF748661	CF748661 UI-M-HJ0-
	15	679.6	15.3	792	12	BM040993	BM040993 603614723
c	16	673.4	15.2	688	12	BG875917	BG875917 IL3-CT021
	17	669.6	15.1	955	12	BM542461	BM542461 AGENCOURT
	18	659.4	14.9	774	14	CA317178	CA317178 UI-M-FW0-
c	19	646	14.6	724	12	BG574669	BG574669 602596672
	20	642.4	14.5	772	14	CB243435	CB243435 UI-CF-FN0
	21	629.4	14.2	801	12	BI685028	BI685028 603310006
c	22	610.4	13.8	665	10	BE538092	BE538092 601062725
	23	609	13.7	648	12	BG721596	BG721596 602695720
	24	608	13.7	734	12	BI656900	BI656900 603284510
c	25	607.4	13.7	638	12	BI667071	BI667071 603291250
	26	603.6	13.6	861	12	BI691674	BI691674 603307363
	27	596.8	13.5	664	14	CD707743	CD707743 EST24270
c	28	596.6	13.4	691	10	BE618504	BE618504 601462944
	29	595.6	13.4	1201	13	BX419085	BX419085 BX419085
	30	584.4	13.2	610	13	BX643048	BX643048 DKFZp781H
c	31	581.4	13.1	607	9	AI590599	AI590599 tw23d07.x
	32	580.8	13.1	1061	12	BG397477	BG397477 602439306
	33	579.4	13.1	960	12	BG756262	BG756262 602713574
c	34	568	12.8	617	14	CD698394	CD698394 EST14917
	35	567.2	12.8	825	14	CF411420	CF411420 CH3#073_E
	36	564.8	12.7	575	10	AW502890	AW502890 UI-HF-BN0
c	37	559.6	12.6	1090	12	BM804157	BM804157 AGENCOURT
	38	558	12.6	610	13	BX475196	BX475196 DKFZp686I
	39	547.4	12.3	693	10	BE535950	BE535950 601062268
c	40	541	12.2	564	13	BX474261	BX474261 DKFZp686J
	41	535.2	12.1	977	10	BE889560	BE889560 601512475
	42	534	12.0	630	9	AI114820	AI114820 HA1429 Hu
c	43	516.8	11.6	523	13	BX106802	BX106802 BX106802
	44	514.8	11.6	531	9	AL120858	AL120858 DKFZp762J
	45	514.2	11.6	747	10	BE963549	BE963549 601657344